Doc Ref. FP25
Appl. No. 10/553,685

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, 5/00, C12P 21/00	A1	**(11) International Publication Number:** WO 89/01036 **(43) International Publication Date:** 9 February 1989 (09.02.89)

(21) International Application Number: PCT/GB88/00602

(22) International Filing Date: 22 July 1988 (22.07.88)

(31) Priority Application Number: 8717430

(32) Priority Date: 23 July 1987 (23.07.87)

(33) Priority Country: GB

(71) Applicant *(for all designated States except US)*: CELLTECH LIMITED [GB/GB]; 216 Bath Road, Slough, Berkshire SL1 4EN (GB).

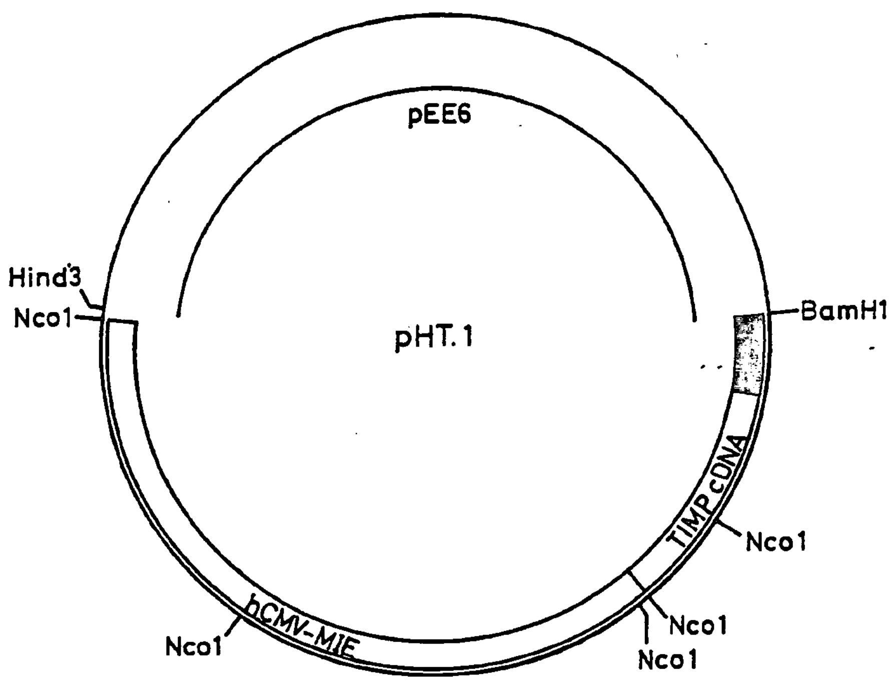
(72) Inventor; and
(75) Inventor/Applicant *(for US only)* : BEBBINGTON, Christopher, Robert [GB/GB]; 20C Trinity Place, Windsor, Berkshire SL4 3AT (GB).

(74) Agent: VOTIER, Sidney, David; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RECOMBINANT DNA EXPRESSION VECTORS



(57) Abstract

The invention provides expression vectors containing the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the major immediate early gene of human cytomegalovirus. Further vectors including the hCMV-MIE DNA linked directly to the coding sequence of a heterologous gene are described. Host cells transfected with the vectors and a process for producing heterologous polypeptides using the vectors and the use of the hCMV-MIE DNA for expression of a heterologous gene are also included within the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		

RECOMBINANT DNA EXPRESSION VECTORS

Field of the Invention

This invention relates to expression vectors containing a DNA sequence from the human cytomegalovirus major immediate early gene, to host cells containing such vectors, to a method of producing a desired polypeptide by using vectors containing said sequence and to the use of said DNA sequence.

Background to the Invention

The main aim of workers in the field of recombinant DNA technology is to achieve as high a level of production as possible of a particular polypeptide. This is particularly true of commercial organisations who wish to exploit the use of recombinant DNA technology to produce polypeptides which naturally are not very abundant.

Generally the application of DNA technology involves the cloning of a gene encoding the desired polypeptide, placing the cloned gene in a suitable expression vector, transfecting a host cell line with the vector, and culturing the transfected cell line to produce the polypeptide. It is almost impossible to predict whether any particular vector or cell line or combination thereof will lead to a useful level of production.

In general, the factors which significantly affect the amount of polypeptide produced by a transfected cell line are: 1. gene copy number, 2. efficiency with which the gene is transcribed and the mRNA translated, 3. the stability of the mRNA and 4. the efficiency of secretion of the protein.

The majority of work directed at increasing expression levels of recombinant polypeptides has focussed on improving transcription initiation mechanisms. As a result the factors affecting efficient translation are much less well understood and defined, and generally it is not possible to predict whether any particular DNA sequences will be of use in obtaining efficient translation.

Attempts at investigating translation have consisted largely of varying the DNA sequence around the consensus translation start signal to determine what effect this has on translation initiation (Kozak M. Cell 41 283-292 (1986)).

Studies involving expression of desired heterologous genes normally use both the coding sequence and at least part of the 5'-untranslated sequence of the heterologous gene such that translation initiation is from the natural sequence of the gene. This approach has been found to be unreliable probably as a result of the 'hybrid nature' of the 5'-untranslated region and the fact that the presence of particular 5-untranslated sequences can lead to poor initiation of translation (Kozak M. Procl. Natl. Acad. Sci. 83 2850-2854 (1986) and Pelletier and Sonenberg Cell 40 515-526 (1985)). This variation in translation has a detrimental effect on the amount of the product produced.

Previous studies (Boshart et al Cell 41 521-530 (1985) and Pasleau et al, Gene 38 227-232 (1985); Stenberg et al, J. Virol 49 (1) 190-199 (1984); Thomsen et al Proc. Natl. Acad. Sci. USA 81 659-663 (1984) and Foecking and Hofstetter Gene 45 101-105 (1986)) have used sequences from the upstream region of the hCMV-MIE gene in expression vectors. These have, however, solely been concerned with the use of the sequences as promoters and/or enhancers. Spaete and Mocarski (J. Virol 56 (1) 135-143, 1985) have used a PstI to PstI fragment of the hCMV-MIE gene encompassing the promoter, enhancer and part of the 5'-untranslated region, as a promoter for expression of heterologous genes. In order to obtain translation the natural 5'-untranslated region of the heterologous gene was used.

In published European Patent Application No. 260148, a method for the continuous production of a heterologous protein is described. The expression vectors constructed contain part of the 5'-untranslated region of the hCMV-MIE gene as a stabilising

sequence. The stabilising sequence is placed in the 5'-untranslated region of the gene encoding the desired heterologous protein i.e. the teaching is again that the natural 5'-untranslated region of the gene is essential for translation.

Summary of the Invention

In a first aspect the invention provides a vector containing a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the major immediate early gene of human cytomegalovirus.

In a preferred embodiment of the first aspect of the invention, the vector includes a restriction site for insertion of a heterologous gene.

The present invention is based on the discovery that vectors containing a DNA sequence comprising the promoter, enhancer and complete 5'-untranslated region of the major immediate early gene of the human cytomegalovirus (hCMV-MIE) upstream of a heterologous gene result in high level expression of the heterologous gene product. In particular, we have unexpectedly found that when the hCMV-MIE derived DNA is linked directly to the coding sequence of the heterologous gene high levels of mRNA translation are achieved. This efficient translation of mRNA is achieved consistently and appears to be independent of the particular heterologous gene being expressed.

In a second aspect the invention provides a vector containing a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the major immediate early gene of human cytomegalovirus upstream of a heterologous gene.

The hCMV-MIE derived DNA according to the second aspect of the invention may be separated from the coding sequence of the

heterologous gene by intervening DNA such as for example by the 5'-untranslated region of the heterologous gene. Advantageously the hCMV-MIE derived DNA may be linked directly to the coding sequence of the heterologous gene.

In a preferred embodiment of the second aspect of the invention, the invention provides a vector containing a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the hCMV-MIE gene linked directly to the DNA coding sequence of the heterologous gene.

Preferably the hCMV-MIE derived sequence includes a sequence identical to the natural hCMV-MIE translation initiation signal. It may however be necessary or convenient to modify the natural translation initiation signal to facilitate linking the coding sequence of the desired polypeptide to the hCMV-MIE sequence, i.e. by introducing a convenient restriction enzyme recognition site. For example the translation initiation site may advantageously be modified to provide an NcoI recognition site.

The heterologous gene may be a gene coding for any eukaryotic polypeptide such as for example a mammalian polypeptide such as an enzyme, e.g. chymosin or gastric lipase; an enzyme inhibitor, e.g. tissue inhibitor of metalloproteinase (TIMP); a hormone, e.g. growth hormone; a lymphokine, e.g. an interferon; a plasminogen activator, e.g. tissue plasminogen activator (tPA) or prourokinase; or a natural, modified or chimeric immunoglobulin or a fragment thereof including chimeric immunoglobulins having dual activity such as antibody-enzyme or antibody-toxin chimeras.

According to a third aspect of the invention there is provided host cells transfected with vectors according to the first or second aspect of the invention.

The host cell may be any eukaryotic cell such as for example plant, or insect cells but is preferably a mammalian cell such as for example CHO cells or cells of myeloid origin e.g. myeloma or hybridoma cells.

In a fourth aspect the invention provides a process for the production of a heterologous polypeptide by culturing a transfected cell according to the third aspect of the invention.

In a fifth aspect the invention provides the use of a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the hCMV-MIE gene for expression a heterologous gene.

In a preferred embodiment of the fifth aspect of the invention the hCMV-MIE derived DNA sequence is linked directly to the DNA coding sequence of the heterologous gene.

Also included within the scope of the invention are plasmids pCMGS, pHT.1 and pEE6hCMV.

Brief Description of the Drawings

The present invention is now described, by way of example only, with reference to the accompanying drawings in which

Figure 1: shows a diagrammatic representation of plasmid pSVLGS.1

Figure 2: shows a diagrammatic representation of plasmid pHT.1

Figure 3: shows a diagrammatic representation of plasmid pCMGS

Figure 4: shows the complete sequence of the promoter-enhancer hCMV-MIE including both the first intron and a modified translation 'start' site

Figure 5: shows a diagrammatic representation of plasmid pEE6.hCMV

Detailed Description of the Embodiments

Example 1

The Pst-1m fragment of hCMV (Boshart et al Cell 41 521-530 (1985) Spaete & Mocarski J. Virol 56 (1) 135-143 (1985)) contains the promoter-enhancer and most of the 5'-untranslated leader of the MIE gene including the first intron. The remainder of the 5'untranslated sequence can be recreated by attaching a small additional sequence of approximately 20 base pairs.

Many eukaryotic genes contain an Ncol restriction site (5'-CCATGG-3') overlapping the translation start site, since this sequence frequently forms part of a preferred translation initiation signal 5'ACCATGPu-3'. The hCMV-MIE gene does not have an Ncol site at the beginning of the protein coding sequence but a single base-pair alteration causes the sequence both to resemble more closely the "Kozak" concensus initiation signal and introduces an Ncol recognition site. Therefore a pair of complementary oligonucleotides were synthesised of the sequence:

```
      GTCACCGTCCTTGACAC
      |||||||||||||||||
  ACGTCAGTGGCAGGAACTGTGGTAC
```

which when fused to the Pst-1m fragment of hCMV will recreate the complete 5'-untranslated sequence of the MIE gene with the single alteration of a G to a C at position -1 relative to the translation initiation codon.

This synthetic DNA fragment was introduced between the hCMV Pst-1m promoter-enhancer leader fragment and a glutamine synthetase (GS) coding sequence by ligation of the Pst-1m fragment and the synthetic oligomer with Ncol digested pSV2.GS to generate a new plasmid, pCMGS (The production of pSV2.GS is described in published International Patent Application No. WO 8704462). pCMGS is shown in Figure 3. pCMGS thus contains a hybrid transcription unit consisting of the following: the synthetic oligomer described above upstream of the

hCMV-MIE promoter-enhancer (where it serves merely as a convenient Pstl - Ncol "adaptor"), the hCMV-MIE promoter and the complete 5' untranslated region of the MIE gene, including the first intron, fused directly to the GS coding sequence at the translation initiation site.

pCMGS was introduced into CHO-KI cells by calcium phosphate mediated transfection and the plasmid was tested for the ability to confer resistance to the GS-inhibitor methionine sulphoximine (MSX). The results of a comparison with pSV2.GS are shown in Table 1.

It is clear that pCMGS can confer resistance to 20 M MSX at a similar frequency to pSV2.GS, demonstrating that active GS enzyme is indeed expressed in this vector.

Table 1

Results of transfection of GS-expression vectors into CHO-KI cells

Vector	no. colonies/10^6 cells resistant to 20μM MSX
pSV2.GS	32
pCMGS	17
Control	0

Example 2

The TIMP cDNA and SV40 polyadenylation signal as used in pTIMP 1 Docherty et al (1985) Nature 318, 66-69, was inserted into pEE6 between the unique HindIII and BamHI sites to create pEE6TIMP. pEE6 is a bacterial vector from which sequences inhibitory to replication in mammalian cells have been removed. It contains the XmnI to BclI portion of pCT54 (Emtage et al 1983 Proc. Natl. Acad. Sci. USA 80,

3671-3675) with a pSP64 (Melton et al 1984: Nucleic Acids. Res. 12, 7035) polylinker inserted in between the HindIII and EcoRI sites. The BamHI and SalI sites have been removed from the polylinker by digestion, filling in with Klenow enzyme and religation. The BclI to BamHI fragment is a 237 bp SV40 early polyadenylation signal (SV40 2770 to 2533). The BamHI to the BglI fragment is derived from pBR328 (375 to 2422) with an additional deletion between the SalI and the AvaI sites (651 to 1425) following the addition of a SalI linker to the AvaI site. The sequence from the BglI to the XmnI site originates from the β-lactamase gene of pSP64.

The 2129 base-pair Ncol fragment containing the hCMV MIE promoter-enhancer and 5' untranslated sequence was isolated from pCMGS by partial Ncol digestion and inserted at the Ncol site overlapping the translation initiation signal of TIMP in pEE6.TIMP to generate the plasmid pHT.1 (shown in Figure 2).

A GS gene was introduced into pHT.1 to allow selection of permanent cell lines by introducing the 5.5K Pvul - BamH1 fragment of pSVLGS.1 (figure 1) at the BamH1 site of pHT.1 after addition of a synthetic BamH1 linker to Pvul digested pSVLGS.1 to form pHT.1GS. In this plasmid the hCMV-TIMP and GS transcription units transcribe in the same orientation.

pHT.1 GS was introduced into CHO-K1 cells by calcium-phosphate mediated transfection and clones resistant to 20μM MSX were isolated 2-3 weeks post-transfection. TIMP secretion rates were determined by testing culture supernatants in a specific two site ELISA, based on a sheep anti TIMP polyclonal antibody as a capture antibody, a mouse TIMP monoclonal as the recognition antibody, binding of the monoclonal being revealed using a sheep anti mouse IgG peroxidase conjugate. Purified natural TIMP was used as a standard for calibration of the assay and all curves were linear in the range of 2 - 20ng ml^{-1}. No non-specific reaction was detectable in CHO-cell conditioned culture media.

One cell line GS.19 was subsequently recloned, and a sub-clone GS 19-12 secretes TIMP at a very high level of 3 x 10^8 molecules/cell/day. Total genomic DNA extracted from this cell line was hybridised with a TIMP probe by Southern blot analysis using standard techniques and shown to contain a single intact copy of the TIMP transcription unit per cell (as well as two re-arranged plasmid bands). This cell line was selected for resistance to higher levels of MSX and in the first selection a pool of cells resistant to 500µM MSX was isolated and recloned . The clone GS-19.6(500)14 secretes 3 x 10^9 molecules TIMP/cell/day. The vector copy-number in this cell line is approx. 20 - 30 copies/cell. Subsequent rounds of selection for further gene amplification did not led to increased TIMP secretion.

Thus it appears that the hCMV-TIMP transcription unit from pHT.1 can be very efficiently expressed in CHO-KI cells at approximately a single copy per cell and a single round of gene amplification leads to secretion rates which are maximal using current methods.

Example 3

In order to test whether the hCMV-MIE promoter-enhancer-leader can be used to direct the efficient expression of other protein sequences, two different but related plasminogen activator coding sequences (designated PA-1 and PA-2) were introduced into CHO-KI cells in vectors in which the protein coding sequences were fused directly to the hCMV sequence.

In both these cases, there is no Ncol site at the beginning of the translated sequence and so synthetic oligonucleotides were used to recreate the authentic coding sequence from suitable restriction sites within the translated region. The sequence of the modified hCMV translation-initiation signal as used in pHT.1 was also built into the synthetic oligonucleotide which then ended in a Pst-1 restriction site. The Pst-1m fragment of hCMV was then inserted at this site to create the complete promoter-enhancer-leader sequence.

The hCMV-plasminogen activator transcription units were introduced into CHO-KI cells after inserting a GS gene at the unique BamH1 site as above and MSX resistant cell lines secreting plasminogen activator were isolated.

The secretion rates of the best initial transfectant cell lines in each case are given in Table 2. From this it is clear that the hCMV promoter-enhancer leader can also be used to direct the efficient expression of these two plasminogen activator proteins.

Table 2

Secretion rates of the different plasminogen activator proteins from transfectant CHO cell lines.

Plasminogen activator	Molecules secreted/cell/day
PA - 1	5.5×10^{7}
PA - 2	1.1×10^{8}

Example 4

pEE6hCMV was made by ligating the Pst-1m fragment of hCMV, HindIII - digested pEE6 and the complementary oligonucleotides of the sequence:

```
    GTCACCGTCCTTGACACGA
    |||||||||||||||||||
ACGTCAGTGGCAGGAACTGTGCTTCGA
```

cDNA encoding an immunoglobulin light-chain was inserted at the EcoRI site of pEE6.hCMV such that the hCMV-MIE promoter-enhancer leader could direct expression of the cDNA and a selectable marker gene containing the SV40 origin of replication was inserted at the BamHI site of each plasmid.

This plasmid was transfected into COS-1 monkey kidney cells by a standard DEAE-dextran transfection procedure and transient expression was monitored 72 hours post transfection. Light chain was secreted into the medium at at least 100ng/ml indicating that light chain can indeed be expressed from a transcription unit containing the entire hCMV-MIE 5'-untranslated sequence up to but not including the translation initiation ATG, followed by 15 bases of natural 5'-untranslated sequence of the mouse immunoglobulin light-chain gene.

CLAIMS

1. A vector containing a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the hCMV-MIE gene.

2 A vector according to claim 1 wherein the vector includes a restriction site for insertion of a heterologous gene.

3. A vector containing a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the hCMV-MIE gene upstream of a heterologous gene.

4. A vector according to Claim 3 wherein the hCMV-MIE DNA is linked directly to the DNA coding sequence of a heterologous gene.

5. A vector according to Claim 4 wherein the hCMV-MIE DNA includes a translation initiation signal.

6. A host cell transfected with a vector according to any of the preceeding claims.

7. A process for the production of a heterologous polypeptide by culturing a host cell according to Claim 6.

8. The use of a DNA sequence comprising the promoter, enhancer and substantially complete 5'-untranslated region including the first intron of the hCMV-MIE gene for expression of a heterologous gene.

9. The use of a DNA sequence according to Claim 8 wherein the hCMV-MIE derived DNA is linked directly to the coding sequence of the heterologous gene.

10. Plasmids pCMGS, pEE6.hCMV and pHT.1

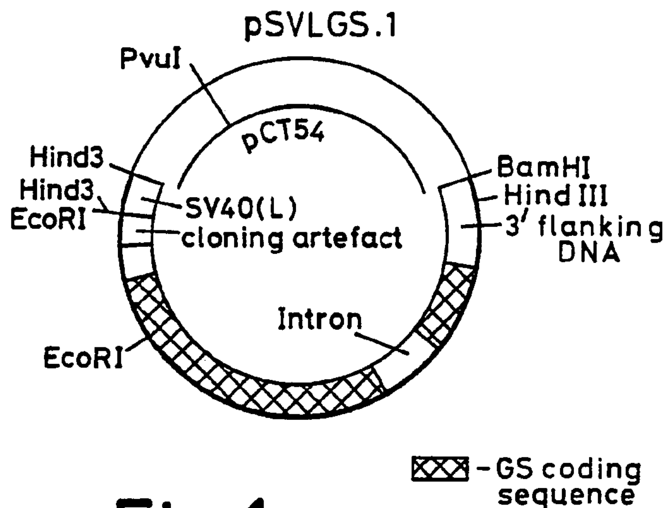

pSVLGS.1
PvuI
pCT54
Hind3
Hind3
EcoRI
SV40(L)
cloning artefact
BamHI
Hind III
3' flanking DNA
Intron
EcoRI
-GS coding sequence

Fig. 1

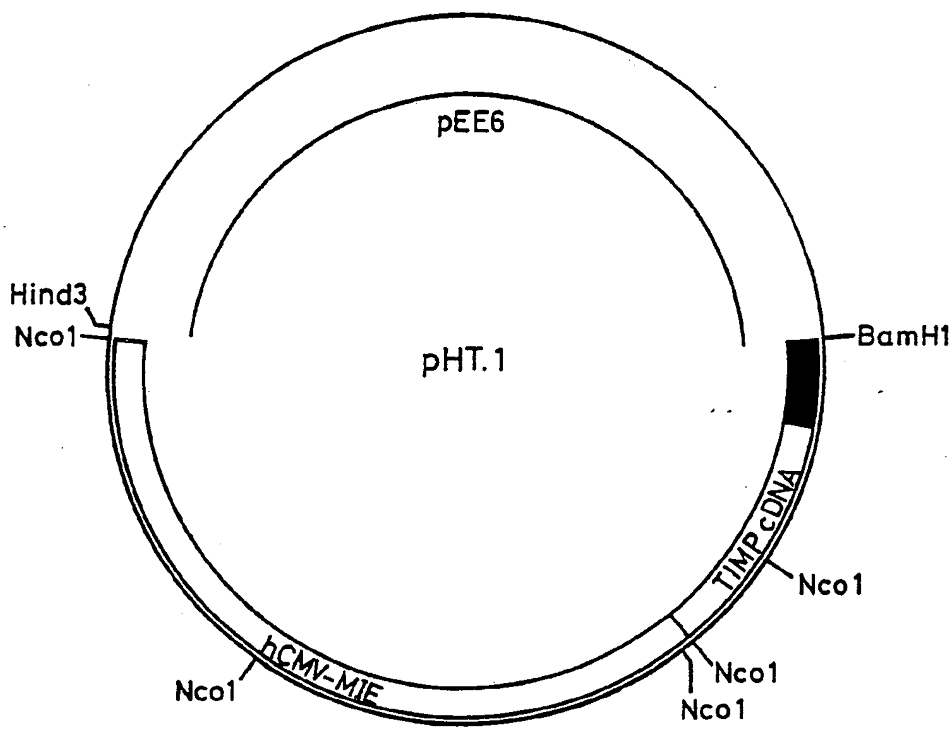

pEE6
pHT.1
Hind3
Nco1
BamH1
TIMP cDNA
Nco1
Nco1
Nco1
hCMV-MIE
Nco1

Fig. 2

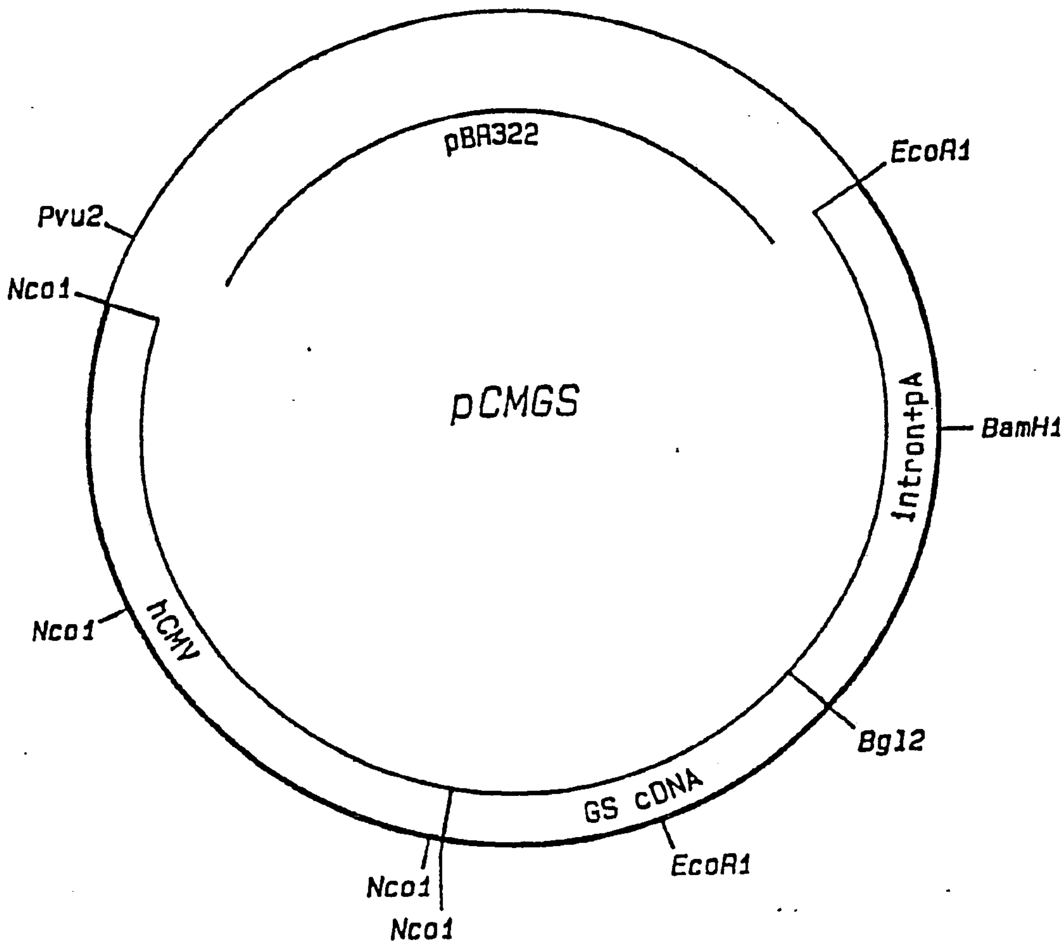

pBR322
EcoR1
Pvu2
Nco1
pCMGS
Intron+pA
BamH1
Nco1
hCMV
Bgl2
GS cDNA
EcoR1
Nco1
Nco1

Fig. 3

```
                                               T
    DNS                          P             t                       AM
    sct                          s             h                       fl
    aoy                          t             3                       1u
    111                          1             2                       31
     //                                                                 /
    CCATGGTGTCAAGGACGGTGACTGCAGTGAATAATAAAATGTGTGTTTGTCCGAAATACG
  1 ---------+---------+---------+---------+---------+---------+ 60
    GGTACCACAGTTCCTGCCACTGACGTCACTTATTATTTTACACACAAACAGGCTTTATGC

    CGTTTTGAGATTTCTGTCGCCGACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCG
 61 ---------+---------+---------+---------+---------+---------+ 120
    GCAAAACTCTAAAGACAGCGGCTGATTTAAGTACAGCGCGCTATCACCACAAATAGCGGC

                            C
                            l
                            a
                            1
    ATAGAGATGGCGATATTGGAAAAATCGATATTTGAAAATATGGCATATTGAAAATGTCGC
121 ---------+---------+---------+---------+---------+---------+ 180
    TATCTCTACCGCTATAACCTTTTTAGCTATAAACTTTTATACCGTATAACTTTTACAGCG

                            E
                            c
                            o
                            R
                            V
    CGATGTGAGTTTCTGTGTAACTGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATAC
181 ---------+---------+---------+---------+---------+---------+ 240
    GCTACACTCAAAGACACATTGACTATAGCGGTAAAAAGGTTTTCACTAAAAACCCGTATG

      E
      c
      o
      R
      V
    GCGATATCTGGCGATAGCGGCTTATATCGTTTACGGGGGATGGCGATAGACGACTTTGGT
241 ---------+---------+---------+---------+---------+---------+ 300
    CGCTATAGACCGCTATCGCCGAATATAGCAAATGCCCCCTACCGCTATCTGCTGAAACCA

    GACTTGGGCGATTCTGTGTGTCGCAAATATCGCAGTTTCGATATAGGTGACAGACGATAT
301 ---------+---------+---------+---------+---------+---------+ 360
    CTGAACCCGCTAAGACACACAGCGTTTATAGCGTCAAAGCTATATCCACTGTCTGCTATA

                                             C BH      N  C
                                             f aa      s  l
                                             r le      i  a
                                             1 11      1  1
                                                /
    GAGGCTATATCGCCGATAGAGGCGACATCAAGCTGGCACATGGCCAATGCATATCGATCT
361 ---------+---------+---------+---------+---------+---------+ 420
    CTCCGATATAGCGGCTATCTCCGCTGTAGTTCGACCGTGTACCGGTTACGTATAGCTAGA
```

Fig. 4A

4/9

```
                  S  C BH
                  s  f aa
                  p  r le
                  1  1 11
                        /
    ATACATTGAATCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCA
421 ---------+---------+---------+---------+---------+---------+ 480
    TATGTAACTTAGTTATAACCGGTAATCGGTATAATAAGTAACCAATATATCGTATTTAGT
     S         C BH
     s         f aa
     p         r le
     1         1 11
                  /
    ATATTGGCTATTGGCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTG
481 ---------+---------+---------+---------+---------+---------+ 540
    TATAACCGATAACCGGTAACGTATGCAACATAGGTATAGTATTATACATGTAAATATAAC

                              H
                              i     M        S
                              n     m        p
                              c     e        e
                              2     1        1
    GCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAAT
541 ---------+---------+---------+---------+---------+---------+ 600
    CGAGTACAGGTTGTAATGGCGGTACAACTGTAACTAATAACTGATCAATAATTATCATTA

    CAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGG
601 ---------+---------+---------+---------+---------+---------+ 660
    GTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTCAAGGCGCAATGTATTGAATGCC

                B                               A  A
                g                               h  a
                1                               a  t
                1                               2  2
    TAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT
661 ---------+---------+---------+---------+---------+---------+ 720
    ATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAACTGCAGTTATTACTGCA

                                         A  A
                                         h  a
                                         a  t
                                         2  2
    ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTAC
721 ---------+---------+---------+---------+---------+---------+ 780
    TACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACTGCAGTTACCCACCTCATAAATG

                  B                     N
                  g                     d
                  1                     e
                  1                     1
    GGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTG
781 ---------+---------+---------+---------+---------+---------+ 840
    CCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGGGGATAAC
```

Fig. 4B

```
    A  A                     B
    h  a                     g
    a  t                     l
    2  2                     1
    ACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACT
841 ---------+---------+---------+---------+---------+---------+ 900
    TGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATACCCTGA

                          S
                          n                   DNS
                          a                   sct
                          B                   aoy
                          1                   111
                                               //
    TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT
901 ---------+---------+---------+---------+---------+---------+ 960
    AAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGATAATGGTACCACTACGCCAAAA

    GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC
961 ---------+---------+---------+---------+---------+---------+ 1020
    CCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGTGCCCCTAAAGGTTCAGAGGTGG

          A  A                 B
          h  a                 a
          a  t                 n
          2  2                 1
    CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC
1021---------+---------+---------+---------+---------+---------+ 1080
    GGTAACTGCAGTTACCCTCAAACAAAACCGTGGTTTTAGTTGCCCTGAAAGGTTTTACAG

    GTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA
1081---------+---------+---------+---------+---------+---------+ 1140
    CATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACATGCCACCCTCCAGATAT

               BH
              BssS                      G   A
              apia                      s   h
              n1Ac                      u   a
              2211                      1   2
               ///
    TAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
1141---------+---------+---------+---------+---------+---------+ 1200
    ATTCGTCTCGAGCAAATCACTTGGCAGTCTAGCGGACCTCTGCGGTAGGTGCGACAAAAC

                                           N
                    B                 D BCGsSX
                    b                 s gfdpam
                    v                 a lriBca
                    2                 1 112223
                                          ////
    ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAA
1201---------+---------+---------+---------+---------+---------+ 1260
    TGGAGGTATCTTCTGTGGCCCTGGCTAGGTCGGAGGCGCCGGCCCTTGCCACGTAACCTT
```

Fig. 4C

```
    CGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACCC
1261---------+---------+---------+---------+---------+---------+ 1320
    GCGCCTAAGGGGCACGGTTCTCACTGCATTCATGGCGGATATCTCAGATATCCGGGTGGG

     B                N
    Ss              N sS
    tt              s pp
    yX              i Hh
    11              1 11
                       /
    CCTTGGCTTCTTATGCATGCTATACTGTTTTTGGCTTGGGGTCTATACACCCCCGCTTCC
1321---------+---------+---------+---------+---------+---------+ 1380
    GGAACCGAAGAATACGTACGATATGACAAAAACCGAACCCCAGATATGTGGGGGCGAAGG

                           E
                           s
                           p
                           1
    TCATGTTATAGGTGATGGTATAGCTTAGCCTATAGGTGTGGGTTATTGACCATTATTGAC
1381---------+---------+---------+---------+---------+---------+ 1440
    AGTACAATATCCACTACCATATCGAATCGGATATCCACACCCAATAACTGGTAATAACTG

                                               P
                                               f
                                               l
                                               M
                                               1
    CACTCCCCTATTGGTGACGATACTTTCCATTACTAATCCATAACATGGCTCTTTGCCACA
1441---------+---------+---------+---------+---------+---------+ 1500
    GTGAGGGGATAACCACTGCTATGAAAGGTAATGATTAGGTATTGTACCGAGAAACGGTGT

                       E
                       c
                       o
                       5
                       7
    ACTCTCTTTATTGGCTATATGCCAATACACTGTCCTTCAGAGACTGACACGGACTCTGTA
1501---------+---------+---------+---------+---------+---------+ 1560
    TGAGAGAAATAACCGATATACGGTTATGTGACAGGAAGTCTCTGACTGTGCCTGAGACAT

                        E
                        c
                        o
                        3
                        1
    TTTTTACAGGATGGGGTCTCATTTATTATTTACAAATTCACATATACAACACCACCGTCC
1561---------+---------+---------+---------+---------+---------+ 1620
    AAAAATGTCCTACCCCAGAGTAAATAATAAATGTTTAAGTGTATATGTTGTGGTGGCAGG

           B
           s                         X                A      A
           p                         h                v      f
           1                         o                a      l
           2                         2                1      3
    CCAGTGCCCGCAGTTTTTATTAAACATAACGTGGGATCTCCACGCGAATCTCGGGTACGT
1621---------+---------+---------+---------+---------+---------+ 1680
    GGTCACGGGCGTCAAAAATAATTTGTATTGCACCCTAGAGGTGCGCTTAGAGCCCATGCA
```

Fig. 4D

7/9

```
       B             B                                   B
       s            Bs                                  Bs
       p            ap                                  ap
       M            n1                                  n1
       2            22                                  22
                     /                                   /
     GTTCCGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCTACATCCGAGCCCTGCTCCC
1681---------+---------+---------+---------+---------+---------+ 1740
     CAAGGCCTGTACCCGAGAAGAGGCCATCGCCGCCTCGAAGATGTAGGCTCGGGACGAGGG

           G                                              H
           s                                              a
           u                                              e
           1                                              1
     ATGCCTCCAGCGACTCATGGTCGCTCGGCAGCTCCTTGCTCCTAACAGTGGAGGCCAGAC
1741---------+---------+---------+---------+---------+---------+ 1800
     TACGGAGGTCGCTGAGTACCAGCGAGCCGTCGAGGAACGAGGATTGTCACCTCCGGTCTG

                                                  D
                                                  s
                                                  a
                                                  1
     TTAGGCACAGCACGATGCCCACCACCACCAGTGTGCCGCACAAGGCCGTGGCGGTAGGGT
1801---------+---------+---------+---------+---------+---------+ 1860
     AATCCGTGTCGTGCTACGGGTGGTGGTGGTCACACGGCGTGTTCCGGCACCGCCATCCCA

                        BH                 N
                      ABsgS                s                 A B
                      vapia                p                 f b
                      an1Ac                B                 l v
                      12211                2                 2 2
                        ///
     ATGTGTCTGAAAATGAGCTCGGGGAGCGGGCTTGCACCGCTGACGCATTTGGAAGACTTA
1861---------+---------+---------+---------+---------+---------+ 1920
     TACACAGACTTTTACTCGAGCCCCTCGCCCGAACGTGGCGACTGCGTAAACCTTCTGAAT

          N                     N
          s                     sP
          p                     pv
          B                     Bu
          2                     22
                                 /
     AGGCAGCGGCAGAAGAAGATGCAGGCAGCTGAGTTGTTGTGTTCTGATAAGAGTCAGAGG
1921---------+---------+---------+---------+---------+---------+ 1980
     TCCGTCGCCGTCTTCTTCTACGTCCGTCGACTCAACAACACAAGACTATTCTCAGTCTCC

                          H
                          iH                             S
                          np                             c
                          ca                             a
                          21                             1
                           /
     TAACTCCCGTTGCGGTGCTGTTAACGGTGGAGGGCAGTGTAGTCTGAGCAGTACTCGTTS
1981---------+---------+---------+---------+---------+---------+ 2040
     ATTGAGGGCAACGCCACGACAATTGCCACCTCCCGTCACATCAGACTCGTCATGAGCAAC
```

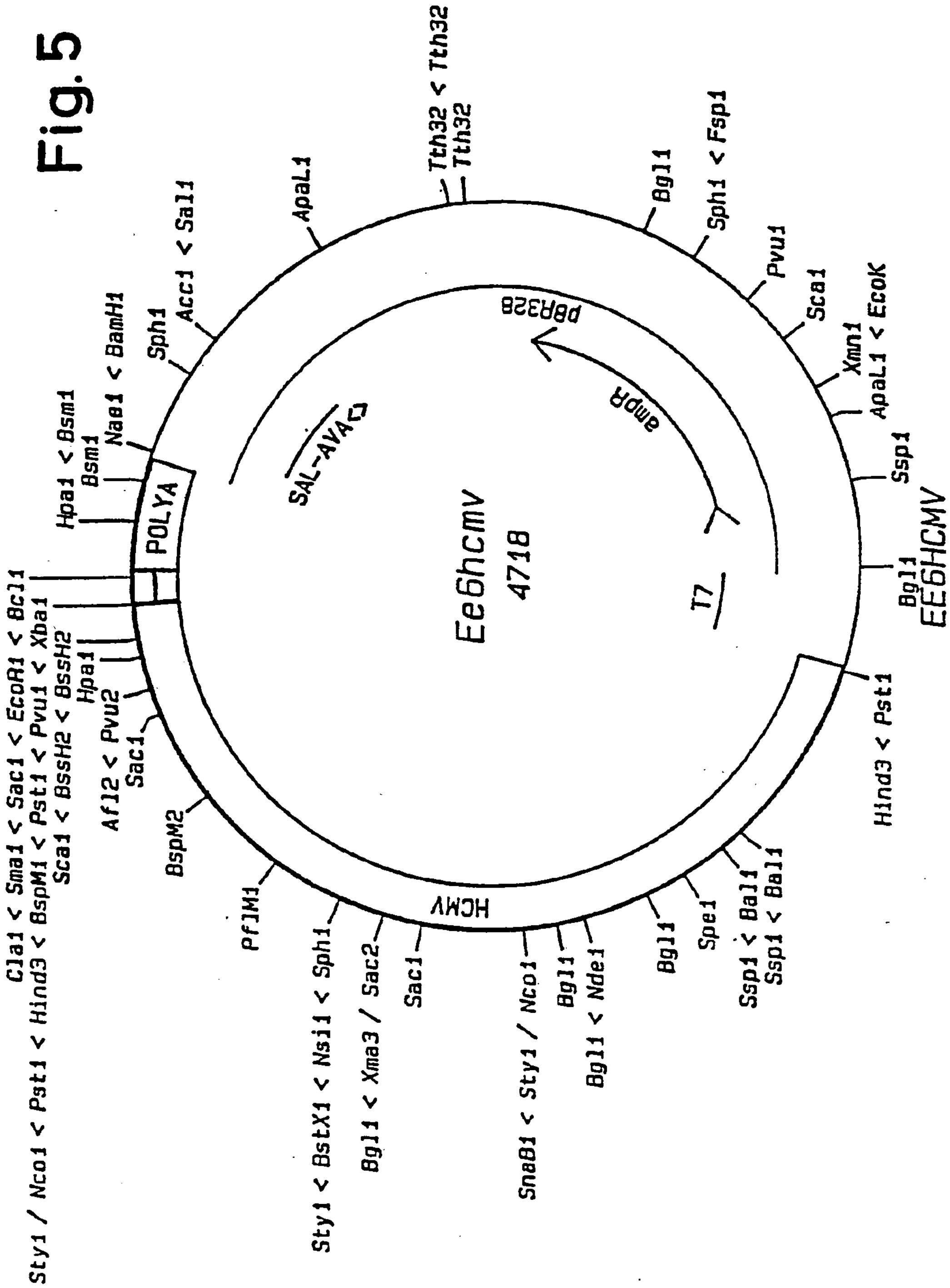
Fig. 4E

```
          B B                                                DNS
          s s                                                sct
          s s                                                aoy
          H H                                                111
          2 2                                                 //
    CTGCCGCGCGCGCCACCAGACATAATAGCTGACAGACTAACAGACTGTTCCTTTCCATGG
2041---------+---------+---------+---------+---------+---------+ 2100
    GACGGCGCGCGCGGTGGTCTGTATTATCGACTGTCTGATTGTCTGACAAGGAAAGGTACC

             P                 DNS
             s                 sct
             t                 aoy
             1                 111
                                //   NcoI
    GTCTTTTCTGCAGTCACCGTCCTTGACACCATG
2101---------+---------+---------+
    CAGAAAAGACGTCAGTGGCAGGAACTGTG
```

Fig. 4F

Fig. 5



INSERT HCMV PROMOTER FRAGMENT IN THE HIND3 SITE OF EE6
Mismatch: 0 MinCuts = 1 MaxCuts: 3

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 88/00602

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC^4: C 12 N 15/00; C 12 N 5/00; C 12 P 21/00

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC^4	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	EP, A, 0173177 (BEHRINGWERKE AG) 5 March 1986 see the whole document --	1-9
X	EP, A, 0173552 (THE UPJOHN CO.) 5 March 1986 see page 9, line 25 - page 10, line 2 --	1-9
X	Gene, vol. 38, no. 1/3, 1985 Elsevier Biomedical Press, (Amsterdam, NL) F. Pasleau et al.: "Growth hormone gene expression in eukaryotic cells directed by the Rous sarcoma virus long terminal repeat or cytomegalovirus immediate-early promoter", pages 227-232; see page 228, second column --	1-9
X	Journal of Virology, vol. 56, no. 1, October 1985 American Society for Microbiology (US) R.R. Spaete et al.: "Regulation of cytomegalovirus gene expression: α and β	1-9

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
26th October 1988	28 NOV 1988
International Searching Authority	Signature of Authorized Officer
EUROPEAN PATENT OFFICE	P.C.G. VAN DER PUTTEN

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category*	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	promoters are trans activated by viral functions in permissive human fibroblasts", pages 135-143 see page 136, last five lines and first paragraph of page 137 (cited in the application) --	
X,P	EP, A, 0260148 (GENENTECH INC.) 16 March 1988, see page 9 (cited in the application)	1-9

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 8800602 SA 23469

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 09/11/88
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0173177	05-03-86	DE-A- 3431140	06-03-86
		AU-A- 4660785	27-02-86
		JP-A- 61111692	29-05-86
EP-A- 0173552	05-03-86	JP-A- 61119192	06-06-86
EP-A- 0260148	16-03-88	FR-A- 2603899	18-03-88
		GB-A- 2197321	18-05-88
		AU-A- 7831787	19-05-88
		JP-A- 63152986	25-06-88
		DE-A- 3730599	07-07-88

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82